Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 9054–9058, December 1987
Evolution

Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs

(plant molecular evolution/molecular clock/mutation rate/organelle DNA/inverted repeat)

KENNETH H. WOLFE*†, WEN-HSIUNG LI*‡, AND PAUL M. SHARP*†

*Center for Demographic and Population Genetics, University of Texas, P.O. Box 20334, Houston, TX 77225; and †Department of Genetics, Trinity College, Dublin 2, Ireland

Communicated by Robert K. Selander, September 8, 1987 (received for review July 7, 1987)

ABSTRACT **Comparison of plant mitochondrial (mt), chloroplast (cp) and nuclear (n) DNA sequences shows that the silent substitution rate in mtDNA is less than one-third that in cpDNA, which in turn evolves only half as fast as plant nDNA. The slower rate in mtDNA than in cpDNA is probably due to a lower mutation rate. Silent substitution rates in plant and mammalian mtDNAs differ by one or two orders of magnitude, whereas the rates in nDNAs may be similar. In cpDNA, the rate of substitution both at synonymous sites and in noncoding sequences in the inverted repeat is greatly reduced in comparison to single-copy sequences. The rate of cpDNA evolution appears to have slowed in some dicot lineages following the monocot/dicot split, and the slowdown is more conspicuous at nonsynonymous sites than at synonymous sites.**

Our current knowledge of the rates and mechanisms of molecular evolution has been derived largely from comparative studies of genes and proteins of animals (1, 2). Only recently has the study of the molecular biology of plants provided sufficient data to allow the evolution of plant genes to be investigated. Since the plant and animal kingdoms diverged about 1000 million years (Myr) ago, their patterns of evolution might have become very different. In fact, plants differ from animals in the organization of their organelle DNA by having a much larger and structurally more variable mitochondrial genome and by having a third (chloroplast) genome (3). So, do the rates of nucleotide substitution differ between animal and plant DNAs? Also, since in mammals mitochondrial DNA (mtDNA) evolves much faster than nuclear DNA (nDNA) (4), do the substitution rates vary greatly among the three plant genomes?

Previous studies based on a few gene sequences or on restriction enzyme mapping have suggested that chloroplast genes have lower rates of nucleotide substitution than mammalian nuclear genes (3, 5) and that plant mtDNA evolves slowly in nucleotide sequence, though it undergoes frequent rearrangement (6). Restriction analysis (3, 7) has also suggested that the large inverted repeat (IR) sequences in chloroplast DNA (cpDNA) have lower rates of nucleotide substitution than the rest of the chloroplast genome. Available DNA sequence data from plants now allow a detailed investigation of the rates of nucleotide substitution in the three plant genomes, reconstruction of the phylogenetic relationships among some higher plants, and comparison of evolutionary rates among lineages.

MATERIALS AND METHODS

DNA sequences were taken from GenBank§ and the literature; the sequences of liverwort and tobacco chloroplast genomes (8, 9) were kindly provided on disk by K. Ohyama and M. Sugiura.

Numbers of nucleotide substitutions in noncoding sequences were calculated by the two-parameter method of Kimura (1); regions in which the correct alignment was not apparent were excluded from the analysis. Protein-coding genes were analyzed by the method of Li *et al.* (10), in which nucleotide substitutions are classified as synonymous (silent) or nonsynonymous (amino acid-changing) and each position in a codon is counted as either a synonymous site, a nonsynonymous site, or one-third synonymous and two-thirds nonsynonymous, depending on the consequences of the substitutions possible at that position. This method provides the numbers of substitutions per synonymous site and per nonsynonymous site (K_S and K_A, respectively), again corrected for multiple hits by Kimura's method. The computer program of Li *et al.* (10) was modified to allow for the differences between the "universal" genetic code and the mitochondrial codes of plants and animals.

In monocot vs. dicot comparisons, wherever more than one sequence is available for a particular gene from monocots or dicots, the values (Table 1) of K (K_S or K_A) and their variances are the means of all possible pairwise comparisons; this procedure tends to overestimate the variance. In pooling different genes to obtain the mean K for each genome, the K value for each gene was weighted by its number of sites (L_S or L_A). The standard error of the mean K was calculated as the square-root of the mean variance

$$\bar{V}_K = \left(\sum_i L_i\right)^{-2} \sum_i L_i^2 \, V_{K_i},$$

where V_{K_i} and L_i are the variance of K and the L_S or L_A for the ith gene.

RESULTS

Rates of Evolution of the Three Plant Genomes. In Table 1 we compare the rates of nucleotide substitution in chloroplast, mitochondrial, and nuclear genes. First, we consider chloroplast and mitochondrial genes. In the comparisons between monocots and dicots the average numbers of nonsynonymous substitutions per site (K_A) in the chloroplast and mitochondrial genomes are similar. In contrast, the average number of synonymous substitutions per site (K_S) in the chloroplast genome is almost 3 times that in the mitochondrial genome, and the ranges of K_S values in large genes

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: mtDNA, mitochondrial DNA; cpDNA, chloroplast DNA; nDNA, nuclear DNA; IR, inverted repeat; SC, single-copy DNA; Myr, million years.

‡To whom reprint requests should be addressed.

§EMBL/GenBank Genetic Sequence Database (1987) GenBank (Bolt, Beranek, and Newman Laboratories, Cambridge, MA), Tape Release 50.0.

Table 1. Numbers of synonymous (K_S) and nonsynonymous (K_A) substitutions per site between species in chloroplast, mitochondrial, and nuclear genes

Gene(s)	Species*	L_S†	$K_S \times 100$	L_A†	$K_A \times 100$
Monocots vs. dicots					
Chloroplast‡					
atpA	MW/PTS	343	59 ± 6	1,166	8 ± 1
atpB	MWBR/PTS	346	66 ± 6	1,138	5 ± 1
atpE	MWBR/PTS	87	59 ± 12	313	18 ± 3
atpF	W/PTS	113	44 ± 8	418	13 ± 2
rbcL	MR/FPTUS	320	72 ± 7	1,102	5 ± 1
psaA	M/PTS	489	55 ± 5	1,743	2 ± 0
psaB	M/PTS	472	50 ± 4	1,724	2 ± 0
psbB	M/TS	342	62 ± 6	1,179	2 ± 0
psbC	M/PTS	328	53 ± 6	1,088	2 ± 0
psbD	M/PTS	239	52 ± 6	817	1 ± 0
psbG	M/TS	161	60 ± 8	567	10 ± 1
petA	WR/VPTES	210	75 ± 10	746	6 ± 1
petB	M/TS	146	64 ± 10	496	1 ± 0
rpS4	M/TS	134	48 ± 8	466	13 ± 2
rpL16§	MI/T	93	53 ± 10	303	8 ± 2
Eight genes¶		355	52 ± 5	1,155	4 ± 1
Total		4177	58 ± 2	14,421	5 ± 0
Mitochondrial					
coxI	GM/YE	355	21 ± 3	1,223	3 ± 1
coxII	MWR/YPE	160	22 ± 4	607	7 ± 1
cob	MW/E	252	9 ± 2	916	3 ± 1
atp9	M/TU	54	28 ± 8	165	2 ± 1
atpA	M/E	333	27 ± 3	1,188	4 ± 1
rpS13	M/T	65	19 ± 6	280	5 ± 1
Total		1219	21 ± 1	4,380	4 ± 0
Nuclear‖					
gapC	B§/D	197	119 ± 16	715	9 ± 1
	B§/T	197	110 ± 14	718	10 ± 1
adh	M1/P	250	191 ± 29	884	11 ± 1
	M2/P	250	>250	884	12 ± 1
	M1/A	248	202 ± 32	886	13 ± 1
	M2/A	247	245 ± 64	887	14 ± 1
Phytochrome	O/Z	724	>250	2,639	24 ± 1
Within monocots (maize vs. wheat)					
Chloroplast					
atpA		342	15 ± 2	1,167	1 ± 0
atpB		343	17 ± 2	1,148	2 ± 0
atpE		88	19 ± 5	321	1 ± 1
atpH		64	8 ± 4	176	0 ± 0
psbH		51	20 ± 7	165	1 ± 1
orf62		45	15 ± 6	138	1 ± 1
Total		934	16 ± 1	3,114	1 ± 0
Mitochondrial					
coxII		163	3 ± 1	614	1 ± 0
cob		250	3 ± 1	911	1 ± 0
Total		413	3 ± 1	1,526	1 ± 0
Within dicots (soybean vs. pea)					
Chloroplast					
psbA		230	23 ± 4	826	0 ± 0
Mitochondrial					
coxII		159	3 ± 1	612	1 ± 0
Within dicots (tobacco vs. petunia)					
Chloroplast					
rbcL		325	8 ± 2	1,104	1 ± 0
psbA		232	5 ± 2	824	0 ± 0
Total		556	7 ± 1	1,928	1 ± 0
Mitochondrial					
atp9		55	2 ± 2	173	0 ± 0

Sequence data can be found in GenBank or the indicated references. Chloroplast: *atpA* (11, 12); *atpBE* (13, 14); *atpF* (12); *atpH* (11, 12); *rbcL* (15–18); *psaAB* (19, 20); *psbA* (21); *psbB* (22); *psbCD* (23); *psbEF* (24); *psbG* and *ndhC* (G. Zurawski, J. Mason, P. Whitfield, personal communication); *psbH* (22, 25, 26); *petA* (27, 28); *petB* (22); *petD* (22, 29); *rpS4* (30); *rpL16* (31); *orf62* (23, 32). Mitochondrial: *coxI* (33, 34); *coxII* (35); *atp9* (36); *atpA* (37). Nuclear: *gapC* (38, 39); *adh* (40); phytochrome (41).

*Species are indicated as monocot/dicot. Species names: A, *Arabidopsis thaliana*; B, barley; D, mustard; E, *Oenothera* species; F, alfalfa; G, sorghum; I, *Spirodela oligorhiza*; M, maize; O, oat; P, pea; R, rice; S, spinach; T, tobacco; U, petunia; V, *Vicia faba*; W, wheat; Y, soybean; Z, zucchini.

†L_S and L_A are the numbers of synonymous and nonsynonymous sites, respectively.

‡Chloroplast genes are named as in refs. 8 and 26; *psbF* refers to the small cytochrome b_{559} gene (39 codons) and *psbH* to the photosystem II 10-kDa phosphoprotein (73 codons). *orf62* is a conserved open reading frame of unknown function (32).

§Partial sequence.

¶Eight genes of <100 codons: *ndhC* (B/PTS) (partial); *atpH* (MW/PTS); *psbE* (W/TES); *psbF* (W/TES); *psbH* (MW/TS); *petD* (M/PTS) (partial); *orf62* (MWB/PTS); *rpL14* (M/T) (partial).

‖The genes for barley cytosolic glyceraldehyde-3-phosphate dehydrogenase (*gapC*) and *Arabidopsis* alcohol dehydrogenase (*adh*) have been shown by blot hybridization of restriction enzyme-digested genomic DNA to be present as single copies per haploid genome. There are probably two *gapC* genes in *Nicotiana tabacum*, consistent with this tobacco species being a recent tetraploid (39, 42). There is no report of the copy number of mustard *gapC*. For *adh* copy-number data, see ref. 40. There are two maize *adh* loci, both of which have been sequenced. Results from two alleles at the *adh1* locus have been averaged. The K_S value for M1 vs. M2 is 103%, so it is probable that the two maize loci arose by duplication after the monocot/dicot split. Pea *adh* is a family of five to eight genes, at least two of which are very closely related. There are at least four phytochrome genes in oat, but possibly only one active gene in zucchini (41).

do not overlap between the two organelle genomes. On a shorter time scale, intrafamily comparisons have also been made, for chloroplast and mitochondrial genes of maize vs. wheat, soybean vs. pea, and tobacco vs. petunia (Table 1). Again, the K_S value is higher in chloroplast genes than in mitochondrial genes, in these cases by a factor of 3–8. This variation in the ratio of the two K_S values could be due to statistical fluctuations (because in some comparisons only one or two genes were used) and/or a real variation in K_S among genes; note that different genes were used in different comparisons. Interpreting the results conservatively, we may conclude that the average synonymous rate in chloroplast genes is about 3 times that in mitochondrial genes.

For the monocot/dicot comparison, nucleotide sequences are available for only two "single-copy" nuclear genes—*gapC* and *adh* (see footnote in Table 1). Since the K_S values for both genes are greater than 100%, it is difficult to obtain reliable estimates of K_S (10); this could be one reason for the large difference in K_S between the two genes. Nevertheless, even the lowest estimate, about 115% for the *gapC* gene, is much greater than the K_S values seen in organelle genes. The nuclear low-copy-number family of genes for phytochrome

shows even higher rates of substitution (Table 1), though this could be an artifact arising from comparison of paralogous loci. Further evidence for a much higher synonymous rate in nuclear genes than in organelle genes is obtained from a comparison of the plastocyanin gene of spinach and *Silene*. The K_S value for this comparison is 126% (Table 2), though spinach and *Silene* are both dicots and have diverged considerably more recently than the monocot/dicot split. Thus the synonymous substitution rate in nuclear genes appears to be at least twice as high as that in chloroplast genes and 5 times higher than that in mitochondrial genes.

In order to consider absolute rates of nucleotide substitution, we must know the divergence times between the taxa compared. Unfortunately, due to the paucity of the plant fossil record, only rough estimates of divergence times are available (Table 2). In particular, the date for the monocot/dicot split could be older than 140 Myr (48, 50), and for this reason the synonymous rates in mitochondrial and chloroplast genes estimated from the monocot/dicot comparison could be overestimates. With this precaution we suggest that the average synonymous substitution rates in plant mitochondrial and chloroplast genes are $0.2–1.0 \times 10^{-9}$ and $1.0–3.0 \times 10^{-9}$, respectively, all rates being expressed in terms of substitutions per site per year. Previous estimates of the synonymous rate in chloroplast genes (5, 11) are somewhat lower than ours, but they were obtained by a method that tends to underestimate synonymous rates, and from much fewer genes. Reliable estimates of the synonymous rate in nuclear genes cannot be made because few genes are available and the K_S values are large (see footnotes in Table 1). Also, the two divergence dates used (Table 2) are uncertain; the spinach/*Silene* date is probably an underestimate because it represents the date by which the pollen of the two organisms had become distinct (45), and the monocot/dicot date may also be too recent, as noted above. Therefore, we can only tentatively suggest that the average synonymous rate in plant nuclear genes is $5.0–30.0 \times 10^{-9}$, probably closer to the lower bound. Hence, this rate may be similar to that in mammalian nuclear genes (44) but could be several times higher (Table 2).

Table 2. Estimated rates of synonymous substitution per 10^9 years in mitochondrial (mt), chloroplast (cp), and nuclear (nuc) genes

Genome	Taxa compared	L_S	$K_S \times 100$	Rate*
Plant				
mt	Maize/wheat	413	3	0.2– 0.3
	Monocot/dicot	1,219	21	0.8– 1.1
cp	Maize/wheat	934	16	1.1– 1.6
	Monocot/dicot	4,177	58	2.1– 2.9
	Angiosp./bryoph.†	10,242	112	1.4– 1.6
nuc	Spinach/*Silene*	123	126	15.8–31.5
	Monocot/dicot	446	161	5.8– 8.1
Primate				
mt	Human/chimpanzee	169	44	21.8–43.7
	Human/orangutan	169	62	19.4–25.9
nuc	Human/chimpanzee	921	2	0.9– 1.9
	Human/orangutan	616	5	1.5– 2.4
Rodent				
mt	Mouse/rat	1,453	109	18.2–54.5
nuc	Mouse/rat	3,886	24	3.9–11.8

The plant K_S values are the mean values from Table 1. The spinach vs. *Silene* comparison is for plastocyanin, which is a single-copy gene (43). The nuclear monocot vs. dicot K_S value is the mean of the *gapC* and *adh1* values. We do not use the maize *adh2* gene because it is about 80% G+C-rich at synonymous codon positions, whereas maize *adh1* is 60%, which is closer to the values for dicot *adh* genes (≈38%). The *gapC*, phytochrome, and plastocyanin genes do not show such great differences in G+C content between species. The rates for primate and rodent nuclear genes are taken from ref. 44. The primate mitochondrial genes are *ndhD* and *ndhF* (both partial sequences) (4). The rodent mitochondrial genes are *coxI, coxII, coxIII, cob, atp6*, and *ndhD* (from GenBank).

*$(K_S \times 10^9)/2T$, where T (divergence time) is 20–40 Myr for spinach vs. *Silene* (45), 50–70 Myr for maize vs. wheat (5, 46), 100–140 Myr for monocots vs. dicots (47, 48), and 350–400 Myr for angiosperms vs. bryophytes (48, 49). The mammalian divergence times are as in ref. 44.

†Angiosperm vs. bryophyte: the complete chloroplast genomes of tobacco and liverwort (refs. 8 and 9; K.H.W., unpublished results).

The above estimate of the synonymous substitution rate in plant mitochondrial genes is roughly 2–5 and 10–20 times lower than that in nuclear genes of primates and rodents, respectively, and 40–100 times lower than that in mammalian mitochondrial genes (Table 2). The mitochondrial/nuclear ratios of the synonymous rates in plants, primates, and rodents are approximately 0.2, 17, and 5, respectively. (The last value may be low due to saturation of transitions in the rodent mitochondrial genes.) The estimated synonymous rate in chloroplast genes is about equal to that in primate nuclear genes and one-quarter of the rodent nuclear rate.

Rate of Evolution of the Chloroplast Inverted Repeat. The outstanding structural feature of the cpDNA genomes of almost all higher plants studied to date is a large IR sequence, varying in length from 10 to 30 kilobases in different species (3). Restriction-mapping studies at the intrafamilial level have suggested that sequence divergence proceeds more slowly in the IR than elsewhere in the chloroplast genome (7, 51). However, due to the low overall sequence divergence seen in these studies, the rate difference could not be accurately quantified, and restriction mapping cannot distinguish between silent and protein-changing substitutions. Our examination of sequence data from different plant families demonstrates that DNA within the IR indeed evolves at a reduced rate (Table 3). It is striking that for silent substitutions the K value is always higher in single-copy (SC) regions than in IR regions. In the spinach (S) vs. tobacco (T) comparison the K values in SC and IR sequences differ by almost 3-fold in noncoding DNA and by 9-fold for silent sites in protein-coding genes. Similar ratios are observed for the soybean (Y) vs. tobacco and *Spirodela* (I) vs. tobacco comparisons. In the latter (monocot vs. dicot) case, the SC noncoding sequences are so divergent that we are unable to align them, whereas IR regions are only ≈8% divergent.

Phylogenetic Relationships and Molecular Clocks. Fifteen chloroplast genes (4776 codons) have been sequenced in three dicots (tobacco, spinach, and pea) and in at least one monocot (usually maize), as well as in liverwort. These dicots represent three different subclasses [Asteridae, Caryophyllidae, and Rosidae, respectively (47)], among which the phylogenetic relationships are not clear. From the pairwise K_A values between these species we have inferred an unrooted phylogenetic tree (Fig. 1) by the neighbor-joining method (59). As expected, the dicots cluster together and the branch leading to the liverwort is long. That all dicots belong to one lineage, and all monocots to another, is further supported by the presence of an intron in the mitochondrial *coxII* gene of monocots but not of dicots (35). Fig. 1 suggests that the dicots diverged quite soon after their split with the monocots and that spinach and tobacco are more closely related to each other than either is to pea. This is in agreement with Ritland and Clegg's (60) recent topology for these species obtained from DNA sequence data for two chloroplast genes, using the unweighted pair-group method, which implicitly assumes rate-constancy.

The phylogenetic tree (Fig. 1) reveals a slowdown in the rate of evolution in the lineages leading to tobacco and spinach. For example, the branch length from node X to the monocots is 2.54%, which is almost twice the distance from this point to tobacco (1.33%). A relative rate test (61) shows

Table 3. Numbers of silent substitutions per site (K) in single-copy (SC) and inverted-repeat (IR) DNA regions of the chloroplast genome

Species pair	SC or IR	DNA region	No. of sites	K^* × 100
		Noncoding DNA		
S/T	SC	*trnTEYD–psbD*	1353	29 ± 2
		atpB–rbcL	600	20 ± 2
		psbH–petB	783	18 ± 2
		Total	2736	24 ± 2
	IR	*trnV*–16S rRNA	1379	8 ± 1
		trnL region	452	10 ± 2
		Total	1831	9 ± 1
I/T	SC	*rpL16* intron	1636	—†
	IR	3′-*rpS12* intron	690	5 ± 1
		23S rRNA–*trnRN*	1306	10 ± 1
		Total	1996	8 ± 1
		Protein-coding genes		
S/T	SC	27 genes‡	5104	37 ± 1
	IR	*rpL2*§	156	4 ± 2
Y/T	SC	*psbA*	230	41 ± 5
	IR	3′-*rpS12*	62	16 ± 5
		rpS7	107	9 ± 3
		rpL2 (partial)	25	9 ± 7
		Total	194	11 ± 3
I/T	SC	*rpL16* (partial)	94	51 ± 10
	IR	3′-*rpS12*	62	3 ± 2

Species names: I, *Spirodela oligorhiza*; S, spinach; T, tobacco (*Nicotiana tabacum*); Y, soybean. Noncoding regions are identified by genes near them, but these genes were not used in the analysis. Sequences are as in Table 1 or GenBank, except for soybean *rpS12* and *rpS7* (52), *Spirodela* 3′-*rpS12* (53), and the spinach *trnTEY* (54), *trnV* (55), and *trnL* (56) regions.

*K = K_S in the case of protein-coding genes.

†Extremely diverged, so that no reliable alignment can be made.

‡These genes are those in Table 1, plus spinach *atpI* and *rpS2* (12), *psbA* (from GenBank), *rpS11* and *rpoA* (57), and *rpS14* (20).

§Sequences are compared only upstream of a one-nucleotide deletion in the spinach (and *Nicotiana debneyi*) sequence. This frameshift causes the carboxyl termini of these proteins to diverge totally from the *N. tabacum*, liverwort, and *Escherichia coli* proteins (9, 8, 58), and hence the aberrant rates of evolution reported for *rpL2* (5).

tobacco to have fewer nonsynonymous substitutions than pea for 11 of the 15 genes studied, when a monocot is used as the reference species (Table 4). Overall, the slowdown in tobacco is highly significant ($P < 10^{-5}$), as is that in the K_A values for spinach ($K_{PM} - K_{SM} = 0.80 \pm 0.18$; $P < 10^{-4}$). However, there is less evidence of a slowdown in the rate of *silent* substitution in the tobacco and spinach lineages—for tobacco there is a significant rate difference ($K_{PM} - K_{TM} = 4.55 \pm 1.86$), but for spinach there is not ($K_{PM} - K_{SM} = 0.06 \pm 1.98$).

Divergence times between species can be estimated by the method of Li and Tanimura (62), which compensates for differences in rates of evolution among lineages. Using the branch lengths in Fig. 1, and taking the monocot/dicot divergence as 100–140 Myr ago (47, 48), we estimate that the branching date for pea is 90–126 Myr ago and the date for the tobacco/spinach split is 81–114 Myr ago. The latter date is somewhat older than the estimate of 70 Myr used by Zurawski and Clegg (5).

DISCUSSION

DNA sequences of higher plants evolve at different rates, depending on whether they are located in the nuclear, chloroplast, or mitochondrial genome. In sharp contrast to the situation in mammals, where mtDNA evolves at least 5 times faster than nDNA, in angiosperms mtDNA evolves at least 5 times more slowly than nuclear sequences (Tables 1 and 2). Transitions make up about 90% of the differences between closely related primate mtDNA sequences (4) but less than 50% of the substitutions in the plant mitochondrial genes studied. Plant and mammalian mitochondrial genomes also differ in that the former frequently undergoes rearrangement and is much larger and more variable in size (3). Therefore, despite containing similar sets of genes, the two mitochondrial genomes clearly evolve in very different ways.

Our analysis suggests that cpDNA evolves at only half the rate of plant nDNA, supporting the view that the chloroplast genome evolves slowly (3, 5). It is, however, less conservative than plant mtDNA because the synonymous rate in chloroplast genes is 3 times higher (Table 1). Since constraints on synonymous codon choice can reduce the rate of synonymous substitution (63), there may be greater constraints on codon usage in the mitochondrion. However, codon usage patterns in chloroplast and mitochondrial genes are very similar in both degree and direction of bias (unpub-

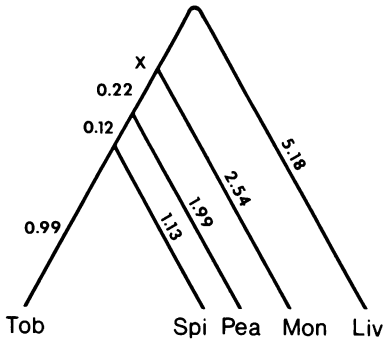

FIG. 1. Phylogenetic tree for tobacco (Tob), spinach (Spi), pea (Pea), monocots (Mon), and liverwort (Liv), reconstructed from nonsynonymous substitution data by the neighbor-joining method (59). The 15 chloroplast genes used are as in Table 4. We did not use synonymous substitutions because the K_S values between liverwort and the other species are too large (>100%) to be useful for this purpose. The pairwise K_A values are as follows: Tob/Spi, 2.12%; Tob/Pea, 3.12%; Tob/Mon, 3.81%; Tob/Liv, 6.51%; Spi/Pea, 3.22%; Spi/Mon, 4.01%; Spi/Liv, 6.70%; Pea-Mon, 4.81%; Pea/Liv, 7.33%; Mon-Liv, 7.72%. Mon = one or more of maize, wheat, barley, and rice; see Table 1.

Table 4. Differences in the number of nonsynonymous substitutions per 100 sites (K_A × 100) between the pea (P) and tobacco (T) lineages, using a monocot species (M) as a reference

Gene	L_A	K_{PT}	K_{PM}	K_{TM}	$K_{PM} - K_{TM}$
atpA	1,160	3.55	8.07	6.40	1.67 ± 0.59*
atpB	1,117	4.00	5.07	4.17	0.90 ± 0.62
atpE	306	11.03	18.37	15.67	2.70 ± 2.23
atpF	401	8.88	15.15	12.05	3.10 ± 1.71
atpH	176	1.30	0.57	0.57	0.00 ± 0.86
rbcL	1,097	3.79	6.02	6.14	−0.12 ± 0.62
psaA	1,716	1.86	3.01	1.88	1.13 ± 0.34†
psaB	1,725	2.93	2.89	1.53	1.36 ± 0.42†
psbC	1,088	0.83	1.96	1.11	0.85 ± 0.28†
psbD	818	0.61	1.61	0.99	0.62 ± 0.28*
psbG	372	1.63	3.16	3.75	−0.59 ± 0.68
petA	747	5.18	6.24	4.60	1.64 ± 0.88
petD	149	0.67	0.68	0.00	0.68 ± 0.68
orf62	139	2.28	5.25	6.03	−0.78 ± 1.37
ndhC	95	5.82	9.05	12.16	−3.11 ± 2.88
Total	11,105	3.12	4.81	3.81	1.00 ± 0.17†

The monocot species used as a reference is maize for all genes except *atpF* (wheat), *petA* (wheat), and *ndhC* (barley) (Table 1).

*Significant at the 5% level.

†Significant at the 1% level.

lished data), suggesting that the constraint on synonymous substitution is similar in the two organelles. This, in turn, suggests that the substitution rates reflect a higher mutation rate in the chloroplast.

Despite a higher synonymous rate in chloroplast genes than in mitochondrial genes, the average nonsynonymous rate is quite similar for the two genomes (Table 1). Thus, while the average K_S/K_A value for the mitochondrial genes is ≈ 5, similar to that for mammalian nuclear genes (10), the ratio for the chloroplast genes is ≈ 11, more than doubled. Interestingly, the homologous chloroplast and mitochondrial *atpA* genes, which encode the α subunit of F_1 ATPase in the respective organelles, both have a K_S/K_A ratio of about 7. Apparently, the other chloroplast genes considered in Table 1 (chiefly components of the photosynthetic apparatus) are, on average, subject to stronger selective constraints than the other mitochondrial genes considered.

A very puzzling finding in this study is that the silent rate in the IR region is at least 3 times lower than that in the rest of the chloroplast genome. This difference in rate has previously been attributed to conservation of the ribosomal RNA genes, which occupy about one-third of the IR region of the chloroplast genome of most higher plants (3), but our results (Table 3) show that other sequences (both noncoding and silent sites in codons) in this region also evolve slowly. If the conservatism of the IR does not reflect a functional constraint, it would imply that the frequency of mutation in this part of the cpDNA molecule is somehow reduced. Alternatively, there may be a bias in the correction of mutations in favor of the original sequences, perhaps connected with the (unknown) mechanism by which the two copies of the IR are maintained as absolutely identical (8, 9).

In plants, as in animals (44, 64), it seems that nucleotide-substitution rates vary among lineages. However, in contrast to the situation in mammals, where the well-documented rate difference between primates and rodents is more pronounced for silent substitutions (61), the most consistent rate change in the plant chloroplast sequences examined is a slowdown in the nonsynonymous rate in some dicots. While such a change in the rate of amino acid replacement may reflect altered selectional constraints on particular proteins, this may not be true in the present case because the rate change is consistent over many genes.

We thank Drs. M. T. Clegg, J. C. Gray, G. S. Hudson, C. J. Leaver, K. Ohyama, M. Sugiura, R. Wu, and G. Zurawski for sending us unpublished DNA sequence data and/or pre-publication manuscripts. We thank Des Higgins and Sue Pagan for their help. This study was supported by National Institutes of Health Grant GM30998.

1. Kimura, M. (1983) *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, Cambridge, U.K.).
2. Li, W.-H., Luo, C.-C. & Wu, C.-I. (1985) in *Molecular Evolutionary Genetics*, ed. MacIntyre, R. J. (Plenum, New York), pp. 1–94.
3. Palmer J. D. (1985) in *Molecular Evolutionary Genetics*, ed. MacIntyre, R. J. (Plenum, New York), pp. 131–240.
4. Brown, W. M., Prager, E. M., Wang, A. & Wilson, A. C. (1982) *J. Mol. Evol.* **18**, 225–239.
5. Zurawski, G. & Clegg, M. T. (1987) *Annu. Rev. Plant Physiol.* **38**, 391–418.
6. Palmer, J. D. & Herbon, L. A. (1987) *Curr. Genet.* **11**, 565–570.
7. Jansen, R. K. & Palmer, J. D. (1987) *Curr. Genet.* **11**, 553–564.
8. Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S.-i., Inokuchi, H. & Ozeki, H. (1986) *Nature (London)* **322**, 572–574.
9. Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Okobata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B. Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H. & Sugiura, M. (1986) *EMBO J.* **5**, 2043–2049.
10. Li, W.-H., Wu, C.-I. & Luo, C.-C. (1985) *Mol. Biol. Evol.* **2**, 150–174.
11. Rodermel, S. R. & Bogorad, L. (1987) *Genetics* **116**, 127–139.
12. Hudson, G. S., Mason, J. G., Holton, T. A., Koller, B., Cox, G. B., Whitfield, P. R. & Bottomley, W. (1987) *J. Mol. Biol.* **196**, 283–298.
13. Moon, E., Kao, T. & Wu, R. (1987) *Nucleic Acids Res.* **15**, 4358–4359.
14. Zurawski, G., Bottomley, W. & Whitfield, P. R. (1986) *Nucleic Acids Res.* **14**, 3974.
15. Moon, E., Kao, T.-H. & Wu, R. (1987) *Nucleic Acids Res.* **15**, 611–630.
16. Aldrich, J., Cherney, B., Merlin, E. & Palmer, J. (1986) *Nucleic Acids Res.* **14**, 9535.
17. Zurawski, G., Bottomley, W. & Whitfield, P. R. (1986) *Nucleic Acids Res.* **14**, 3975.
18. Aldrich, J., Cherney, B., Merlin, E. & Palmer, J. (1986) *Nucleic Acids Res.* **14**, 9534.
19. Lehmbeck, J., Rasmussen, O. F., Bookjans, G. B., Jepsen, B. R., Stummann, B. M. & Henningsen, K. W. (1986) *Plant Mol. Biol.* **7**, 3–10.
20. Kirsch, W., Seyer, P. & Herrmann, R. G. (1986) *Curr. Genet.* **10**, 843–855.
21. Aldrich, J., Cherney, B., Merlin, E., Christopherson, L. & Williams, C. (1986) *Nucleic Acids Res.* **14**, 9536.
22. Rock, C. D., Barkan, A. & Taylor, W. C. (1987) *Curr. Genet.* **12**, 69–77.
23. Bookjans, G., Stummann, B. M., Rasmussen, O. F. & Henningsen, K. W. (1986) *Plant Mol. Biol.* **6**, 359–366.
24. Carrillo, N., Seyer, P., Tyagi, A. & Herrmann, R. G. (1986) *Curr. Genet.* **10**, 619–624.
25. Hird, S. M., Dyer, T. A. & Gray, J. C. (1986) *FEBS Lett.* **209**, 181–186.
26. Westhoff, P., Farchaus, J. W. & Herrmann, R. G. (1986) *Curr. Genet.* **11**, 165–169.
27. Wu, N., Cote, J. C. & Wu, R. (1986) *Gene* **50**, 271–278.
28. Ko, K. & Straus, N. A. (1987) *Nucleic Acids Res.* **15**, 2391.
29. Phillips, A. L. & Gray, J. C. (1984) *Mol. Gen. Genet.* **194**, 477–484.
30. Ben Tahar, S., Bottomley, W. & Whitfield, P. R. (1986) *Plant Mol. Biol.* **7**, 63–70.
31. Posno, M., van Vliet, A. & Groot, G. S. P. (1986) *Nucleic Acids Res.* **14**, 3181–3195.
32. Quigley, F. & Weil, J. H. (1985) *Curr. Genet.* **9**, 495–503.
33. Grabau, E. A. (1986) *Plant Mol. Biol.* **7**, 377–384.
34. Hiesel, R., Schobel, W., Schuster, W. & Brennicke, A. (1987) *EMBO J.* **6**, 29–34.
35. Grabau, E. A. (1987) *Curr. Genet.* **11**, 287–293.
36. Young, E. G., Hanson, M. R. & Dierks, P. M. (1986) *Nucleic Acids Res.* **14**, 7995–8006.
37. Schuster, W. & Brennicke, A. (1986) *Mol. Gen. Genet.* **204**, 29–35.
38. Chojecki, J. (1986) *Carlsberg Res. Commun.* **51**, 203–210.
39. Shih, M.-C., Lazar, G. & Goodman, H. M. (1986) *Cell* **47**, 73–80.
40. Llewellyn, D. J., Finnegan, E. J., Ellis, J. G., Dennis, E. S. & Peacock, W. J. (1987) *J. Mol. Biol.* **195**, 115–123.
41. Lissemore, J. L., Colbert, J. T. & Quail, P. H. (1987) *Plant Mol. Biol.* **8**, 485–496.
42. Okamuro, J. K. & Goldberg, R. B. (1985) *Mol. Gen. Genet.* **198**, 290–298.
43. Rother, C., Jansen, T., Tyagi, A. & Herrmann, R. G. (1986) *Curr. Genet.* **11**, 171–176.
44. Li, W.-H., Tanimura, M. & Sharp, P. M. (1987) *J. Mol. Evol.* **25**, 330–342.
45. Muller, J. (1981) *Bot. Rev.* **47**, 1–142.
46. Chao, S., Sederoff, R. & Levings, C. S., III (1984) *Nucleic Acids Res.* **12**, 6629–6644.
47. Cronquist, A. (1981) *An Integrated System of Classification of Flowering Plants* (Columbia Univ. Press, New York).
48. Stewart, W. N. (1983) *Paleobotany and the Evolution of Plants* (Cambridge Univ. Press, Cambridge, U.K.).
49. Gray, J. (1986) *Nature (London)* **322**, 501–502.
50. Giannasi, D. E. & Crawford, D. J. (1986) *Evol. Biol.* **20**, 25–248.
51. Clegg, M. T., Rawson, J. R. Y. & Thomas, K. (1984) *Genetics* **106**, 449–461.
52. von Allmen, J.-M. & Stutz, E. (1987) *Nucleic Acids Res.* **15**, 2387.
53. Posno, M., Verweij, W. R., Dekker, I. C., de Waard, P. M. & Groot, G. S. P. (1986) *Curr. Genet.* **11**, 25–34.
54. Holschuh, K., Bottomley, W. & Whitfield, P. R. (1984) *Plant Mol. Biol.* **3**, 313–317.
55. Briat, J. F. & Dalmon, J. (1986) *Nucleic Acids Res.* **14**, 8223.
56. Zhou, D.-X., Quigley, F. & Mache, R. (1987) *Nucleic Acids Res.* **15**, 3621.
57. Sijben-Muller, G., Hallick, R. B., Alt, J., Westhoff, P. & Hermann, R. G. (1986) *Nucleic Acids Res.* **14**, 1029–1044.
58. Zurawski, G. & Zurawski, S. M. (1985) *Nucleic Acids Res.* **13**, 4521–4526.
59. Saitou, N. & Nei, M. (1987) *Mol. Biol. Evol.* **4**, 406–425.
60. Ritland, K. & Clegg, M. T. (1987) *Am. Nat.* **130**, S74–S100.
61. Wu, C.-I. & Li, W.-H. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 1741–1745.
62. Li, W.-H. & Tanimura, M. (1987) *Nature (London)* **326**, 93–96.
63. Sharp, P. M. & Li, W.-H. (1987) *Mol. Biol. Evol.* **4**, 222–230.
64. Britten, R. J. (1986) *Science* **231**, 1393–1398.